Regional convergence and agglomeration in Argentina: a spatial panel data approach

Nicole Madariaga, Sylvie Montout, Patrice Ollivaud

UNIVERSITÉ PARIS 1 – PANTHÉON - SORBONNE –

TEAM
Théorie Et Applications en Microéconomie et Macroéconomie
Université Paris I Panthéon Sorbonne
UMR CNRS

Cahiers de la MSE

Regional convergence and agglomeration in Argentina : a spatial panel data approach

Nicole MADARIAGA, TEAM

Sylvie MONTOUT, TEAM

Patrice OLLIVAUD, OECD

2005.06

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13

http://mse.univ-paris1.fr/Publicat.htm

ISSN : 1624-0340

Regional Convergence and Agglomeration in Argentina: a spatial panel data approach*

Nicole Madariaga† Sylvie Montout‡ and Patrice Ollivaud§

February 1, 2005

*We are very grateful to the CEPAL Office in Buenos Aires, to Ricardo Martinez who provided us Argentina's data and to ProvInfo department from the Secretaría de Provincias in Buenos Aires.

†TEAM-CNRS, Université Paris 1 Panthéon Sorbonne. Corresponding author:nicole.madariaga@malix.univ-paris1.fr

‡TEAM-CNRS, Université Paris 1 Panthéon Sorbonne.

§OECD, Economic Department. The views expressed in this paper are those of the authors and do not necessarily reflect those of the OECD or of the governments of its member countries.

Résumé

Cet article étudie la convergence des revenus par tte entre les 23 provinces argentines de 1983 à 2002. L'objet de ce papier est d'appliquer de nouvelles méthodes d'estimation en suivant la procédure en deux étapes de Badinger *et al.* (2004). Nous combinons une méthode de filtrage spatial des variables éliminant l'autocorrélation spatiale (Getis et Griffith, 2002) et un estimateur adapté à un panel dynamique (en utilisant un estimateur MMG en différences premires et en systme). Nos estimations sur données filtrées concluent à une convergence conditionnelle entre les provinces argentines et à un impact significatif de la variable agglomération sur le taux de croissance. Ainsi, nos résultats montrent qu'ignorer les distorsions spatiales dues à la proximité géographique à des estimations erronées et sous-estiment la vitesse de convergence notamment des provinces éloignées de Buenos Aires. De plus, l'estimation des effets d'agglomération est améliorée lorsque l'autocrrélaion spatiale est éliminée.

mots clés: croissance, convergence, dépendance spatiale, filtrage spatial, panels dynamiques, Méthode des Moments généralisés

Abstract

This paper examines the per capita income convergence process among 23 Argentinian provinces over the period 1983-2002. The purpose of this work is to apply new estimation methods following two-step procedure as in Badinger *et al.* (2004). We combine a spatial filtering of variables to remove the spatial correlation (Getis and Griffith, 2002) and suitable estimators for dynamic panels (using first-differenced and system GMM estimators). Our estimations on filtered variables reveal a conditional convergence process between Argentinian provinces and a positive and significant impact of agglomeration variables on growth rate. Hence, our results show that ignoring spatial distortions due to geographic proximity misleads estimations and underestimates the speed of convergence specifically for provinces which are distant from Buenos Aires. Moreover, we improve estimations of agglomeration effects when spatial autocorrelation is removed.

Key words: growth, convergence, spatial dependence, spatial filtering, dynamic panels, Generalized Method of Moments

JEL-classification: C23, O00, R11

1 Introduction

Several empirical works investigated the issue of whether trade liberalization induces a convergence process of income levels within different areas of the world. Most of them concluded to a positive relation[1] despite Rodriguez and Rodrik (1999) cast doubt on the robustness of these results. However, if trade openness seems to enhance the convergence between countries, its impact on internal regional convergence is not so clear. In Madariaga, Montout and Ollivaud (2004), the emergence of the MERCOSUR since 1991 have demonstrated that per capita income convergence is not merely a result of trade integration, but that prior macroeconomic stability and growth are crucial conditions for regional convergence.

Nowadays an ambitious project implementing the hemispheric Free Trade Area of the Americas (FTAA)[2] represents a challenge for the MERCOSUR trading area. Since the election of presidents Luiz Incio Lula da Silva and Néstor Kirchner, Argentina and Brazil have been wishing to emphasize the MERCOSUR cohesion to reinforce their economic and political force facing the United States (US). In such a context, the negotiation force of Argentina to carry on its harmonious cooperation with Brazil partly depends on its economic performances and its internal cohesion. These latter have been strongly eroded since 2001 crisis. In particular, the internal consistency of income per capita distribution is of a major concern: indeed, strong income and agglomeration inequalities characterize the country.

In spite of stagnation and recession episodes suffered by Argentina during last decades, this country has a real value of gross domestic product per inhabitant far over Latin America average. Even so, variables in favour of economic growth and convergence like public educational expenditures and industrial investments remain relatively low in Argentina according to international standards of industrial countries and other developing areas like South-East Asia. Other variables related to economic development, such as foreign trade, also evolve very slowly in comparison with other developing areas that have already reached higher levels of development.

Porto (1994, 1996), Marina (2001), Willington (1998) and Utrera and Koroch (1998) studied regional convergence within Argentina applying the "traditional" cross-sectional approach. Their main results clearly give evidence against the existence of absolute beta (β)-convergence and sigma (σ)-convergence but, in some cases, in favour of the conditional β-convergence. One of the most important implications brought about the neoclassical growth theory is the conditional convergence; that is, poor provinces face growth rates above rich ones *conditionally* to their structural characteristics defining their own steady state (Barro and Sala-i-Martin, 1995). Hence, empirical studies show that Argentinean provinces present diverse structural features. Because of this, even though they seem to convergence toward their respective steady states, they may not converge among themselves. At least, part of structural differences are explained by different education levels and public policies.

As pointed above, the convergence is a crucial issue in Argentina because of great regional inequalities and the prominent role of the capital, Buenos Aires, in production activities. An important share of industrial activities is localized in Buenos Aires, which represents the largest share of national GDP (above 60%), population (more than 30%) and employment (more than 70%) in 2001. As a consequence, this article tests whether the presence of strong disparities in activities distribution induces structural differences explaining conditional convergence.

The new economic geography and growth theories have recently shown the existence of interactions between agglomeration of activities and growth (Baldwin *et al.*, 2003). These theories insist on the central role of externalities geographically concentrated played on spatial concentration and growth mechanisms. Hence, there is a strong relation between the *core-periphery* concentration of the activities distribution and the unequal development of regions.

[1]Edwards (1998), for example, provides a literature revue on this relation.

[2]This project includes all the countries of the western hemisphere except Cuba.

Traditional approaches assume each region as an isolated entity. In this case, the role of spatial dependence is completely neglected when it is an important force in the process of convergence (Rey and Montouri, 1999). Spatial autocorrelation, referring to the absence of independence between geographic observations, is defined as the correlation of a variable with itself proceeding from the geographic distribution of data. Incomes may be correlated with those of neighbouring provinces or regions. As a result, the growth of Argentinian provinces may be characterized by a spatial autocorrelation that have to be taken into account to specify correctly the econometric model[3]. In fact, works using a spatial econometric framework to investigate regional convergence in a cross-sectional analysis are very recent in the literature. Abreu *et al.* (2004) proposes an exhaustive survey of empirical studies on convergence addressing the importance of spatial factors. They underline that 68% of studies applying spatial econometrics analysis cover European regions and only 6% to Latin America. However, none of them is devoted to the Argentinian case.

As we will show, the regional growth in Argentina is characterized by spatial dependence that must be taken into account in a correctly specified convergence model. However, spatial econometric analysis is to a great extent constrained to cross sections[4] (Rey and Montouri (1999), or Baumont *et al.* (2000)) and static panels but no dynamic elements are taken into account. This is why this article is devoted to the application of recent econometric methods estimating dynamics panels addressing the issue of spatial dependence problem.

Furthermore, two sources of inconsistency are distinguished in cross-sectional empirical works on growth. First, the incorrect treatment of *country-specific effects*, represented by differences in technology or tastes, gives rise to omitted variable bias[5]. As a result, Islam (1995) proposes the Within estimator for panel data that allows to control for country-specific effects. However, this estimator considers the explanatory set as perfectly exogenous. Second, a strong theoretical argument reveals that at least a subset of explanatory variables should be expected to be endogenous. We propose to solve these problems by dealing with dynamic panel data models using first-differenced (Caselli *et al.* (1996), Henderson (2000), Levine *et al.* (2000), Easterly *et al.* (1997)) and system Generalized Method of Moments (GMM) estimator (Bond *et al.* (2001), Yudong and Weeks (2000) or Badinger *et al.* (2004)). Studies covering Latin America and specifically Argentina using the GMM estimator to deal with dynamic panel data are very few. To our knowledge, only Estearly *et al.* (1997) employ first-differenced GMM to test the convergence process of Latin America.

The aim of this paper is to apply a new method of estimating convergence that combines spatial filtering and dynamic panel data econometrics. As a consequence, we adopt the methodology used by Badinger *et al.* (2004) in order to estimate a dynamic panel data convergence. We employ a two-step procedure. The first step applies a filtering technique as proposed in Getis and Griffith (2002) to remove spatial autocorrelation from data. The second one estimates the speed of convergence by applying GMM estimators. The augmented Solow model is used to provide the main theoretical framework. Later on, we add to the growth equation estimation agglomeration variables measuring the concentration of employment in the manufacturing sector. As in Madariaga, Montout and Ollivaud (2004), we compute location quotients, Z_iPOP et Z_iAREA, to control for the impact of the concentration of production activities on growth of Argentina provinces. Hence, we check if strong agglomeration disparities may explain different economic performances between provinces, once the spatial autocorrelation is removed.

[3]From an econometric point of view, the spatial dependence between observation does not biases the OLS estimator but estimations are not robust because of spatial autocorrelation in the error term (Beaumont *et al.*, 2000).

[4]For a summary of the literature see Badinger *et al.* (2004).

[5]The initial technological level of a country is correlated with an independent variable, the initial level of per capita income.

The paper proceeds as follows: section 2 presents the spatial filtering technique describing the spatial autocorrelation test and the filtering method of variables removing spatial autocorrelation. The 3 section provides information on different estimators to deal with dynamic panel data and the results of estimations. Section 4 analyzes the link between regional convergence and geography. Finally, section 5 summarizes and concludes.

2 The Spatial filtering

Many studies show the importance of spatial patterns in convergence studies: Argentinian provinces should not be treated as isolated economies (Fingleton (1999) and Rey and Montouri (1999)). It should be assumed that the per capita production growth of a province is linked to its neighbourhood. Upton and Fingleton (1985) affirm that spatial autocorrelation exists whenever there is systematic spatial variation in values across a map. Spatial autocorrelation coefficient measures both the proximity and the similarity of Argentinian province characteristics:

(1) If the coefficient is positive, neighbouring provinces are more similar than distant provinces.

(2) If the coefficient is negative, neighbouring provinces are more dissimilar than distant provinces.

(3) Finally, without any spatial autocorrelation, there is no relation between province proximity and their degree of similarity.

Concepts of "proximity" and "similarity" are evaluated thanks to a spatial weight matrix. Spatial autocorrelation may be due to spillovers phenomenon (like technology spreadings) from one or several provinces implying that the frequency or the intensity of the spreading phenomenon depends on the distance from provinces of origin. The interaction process like goods, workers, capital flows or spatial externalities may also cause spatial autocorrelation. In the context of our dataset, the spatial autocorrelation measures the degree to which the production of a province in Argentina is similar to the production of neighbouring provinces.

A central issue in this paper is whether spatial interaction is important with respect to regional convergence. Then, we firstly proceed to a Moran's I autocorrelation test for spatial dependence and secondly use spatial filtering techniques by computing G_i statistics (Getis and Ord, 1992). This filter allows us to estimate the dataset with GMM methods and to respect the assumption of spatially uncorrelated errors.

2.1 The spatial weight matrix

If provinces are interdependent, the dimension, the structure and the relative position of the province for each variable (per capita production, investment, etc) have to be considered in relation to other provinces. We evaluate geographical connection thanks to an exogenous spatial weight matrix W. This matrix is used to evaluate the covariance of variables associated to each location. It contains information about the relative dependence between n regions. Its elements w_{ij} are derived from the distance between region i and j capitals. We suppose that the intensity of the relation between provinces depends on the distance between capitals. Various indicators can be used according to the expected influence of distance d_{ij}. As Badinger *et al.* (2004), we adopt the most used negative exponential function form:

$$w_{ij} = exp(-d_{ij}.\beta_E)$$

where $0 < \beta_E < \infty$ is an exogenous distance decay parameter. This functional form seems to be perfectly fit the Argentinian case as the quality of infrastructures makes the distance being a major obstacle to production and technologies movement. To make easy the interpretation of spatial autocorrelation, the spatial weight matrix is row-normalized, that is weighting elements w_{ij} are divided by the sum of weighting for the corresponding row. Hence, the normalized weighting may be interpreted as the province share in the total spatial effect of the country.

Table 1: Spatial autocorrelation of income 1983-2002

	Y	X		
		$n+g+\delta$	s_k	s_{h2}
$I_{83}(z_I)$	0,26 (8,82)***	0,23 (7,57)***	0,07 (2,50)**	0,03 (2,3)**
$I_{88}(z_I)$	0,26 (8,98)***	0,23 (7,57)***	0,06 (2,13)**	0,08 (2,3)**
$I_{93}(z_I)$	0,20 (5,41)***	0,15 (4,05)***	0,05 (2,16)**	0,11 (2,91)***
$I_{98}(z_I)$	0,19 (5,08)***	0,15 (4,05)***	0,06 (2,16)**	0,20 (4,77)***
$I_{02}(z_I)$	0,15 (4,04)***	0,15 (3,87)***	0,11 (3,20)***	0,21 (4,73)***

** significant at 5%; *** significant at 1%. One-tail test.
Test results only account for 22 regions (E(I)=-0.047),
except in 1983 (14 regions and E(I)=0,077) and 2002 (21 regions and E(I)=0,05).

2.2 Global spatial autocorrelation measure

The most widely used global measure of spatial autocorrelation is Moran's I (Moran 1950). This statistic tests for spatial patterns that differ from randomness. The Moran's I is given by:

$$I = \frac{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N} w_{ij} Z_i Z_j}{S_0} \Big/ \frac{\sum\limits_{i=1}^{N} Z_i^2}{N}$$

where $S_0 = \sum\limits_i \sum\limits_j w_{ij}$. w_{ij} corresponds to the weighting of the distance between provinces i and j in the matrix W. $Z_i = y_i - \bar{y}$, where $y_i$is the value of the per capita production (or the investment, etc) observed in province i and $\bar{y}$ is its mean. The numerator corresponds to the covariance between two provinces, weighted by a parameter, w_{ij}/S_0, related to the distance between them. The covariance is normalized by the denominator, which is the total observed variance.

Under the null hypothesis of no global spatial autocorrelation, the expected value of I is given by $E(I) = -1/(N-1)$. If I is larger (smaller) than $E(I)$, the distribution of the variable Y is considered positively (negatively) autocorrelated, meaning that the value taken by Y at each location i tends to be similar (dissimilar) to values taken on by Y at spatially contiguous locations. Inference is based on z-values given by $z_I = \frac{I-E(I)}{\sigma(I)}$ where $\sigma(I)$ is the standard deviation of I. We can then test the normal distribution of z_i and check its significance by means of a standard normal table.

Table 1 shows the results of the Moran's I test on our dependent variable, the per capita GDP (y), the growth rate of population corrected for the rate of technological progress and for the depreciation of capital[6], $n+g+\delta$, s_k the share of physical investment, and s_h the share of human capital investment, performed on 5 cross-sections with 22 Argentinean provinces[7]. Results always evidence a significant spatial autocorrelation. The standard normally distributed Moran's I values range from 4,04 to 8,98 with income Y and from 2,10 to 7,57 with other variables[8]. We then compute filtering techniques on the dataset before estimating growth equations.

[6]The sum of these two parameter is equal to 0,5 as in MRW.

[7]Argentina counts for 23 provinces but data on secondary school of the province of Río Negro are never available, then the Moran's I only considers 22 provinces. From 1983 to 2002, the number of provinces accounted is reduced because of other lacks in the dataset.

[8]Tests on agglomeration variables confirm the presence of a positive spatial autocorrelation.

2.3 The filtering method

Filtering methods allow to convert spatially autocorrelated variables into spatially independent variables. The filtering of data removes geographical disparities related to the spatial autocorrelation that could underestimate the convergence effect between Argentinian provinces. In Argentina, the central role played by the province of Buenos Aires on other provinces is particularly emphasized. The importance of geographical links may underestimate the capacity of distant provinces to reduce the gap with per capita income of Buenos Aires. The filtering may become essential to the convergence test.

Getis and Griffith (2002) propose a method inspired on Getis and Ord (1992) that computes the G_i statistics to identify spatial agglomerative patterns with high-value clusters or low-value clusters. This method is chosen because its application is more intuitive and its calculation is easier than other filtering methods. hence, we remove spatial autocorrelation from data before conducting GMM regression analysis.

The G_i depends on the selected weight matrix and is defined as a weighted and normalized average of observations from a relevant variable:

$$G_i = \sum_j w_{ij} X_j / \sum_j X_j, \qquad i \neq j$$

where X_j is the variable to be filtered.

The realization of G_i when no spatial autocorrelation occurs is according to Badinger *et al.* (2004):

$$E\left[G_i\right] = \sum_j w_{ij} / \left(N-1\right)$$

Thus, the ratio between this last value and G_i will indicate the importance of the spatial dependence. The filtered observation is then:

$$X_i^s = X_i \left[\sum_j w_{ij} / \left(N-1\right)\right] / G_i$$

This method has been applied to all series in *levels* as indicated in Getis and Griffith (2002)[9] with the spatial weights matrix discussed above. We construct the variable from initial filtered series. An important point is the choice of the distance decay parameter. Following Niebuhr (2003), a transformed decay parameter γ_E $(0 \leq \gamma_E \leq 1)$ is obtained with $\gamma_E = 1 - e^{-\beta_E \cdot D_{MEAN}}$ where D_{MEAN} denotes the average distance between centers of immediately neighbouring regions. Thus, γ_E measures the percentage decrease of spatial effects if distance expands by a given unit.

It is assumed that spatial interactions such as commuting, migration or interregional trade are exposed to the frictional effects of geographical distance. These geographical impediments gain in strength with increasing γ_E, so that the spatial effects decline with the distance. In order to check the sensitivity of results with respect to a variation of W, the whole range of γ_E is considered throughout the analysis.

3 Estimators for dynamic panel data models

Following Mankiew, Romer and Weil (1992) (MRW) we may write an autoregressive form of the growth model as

[9] The filtering technique should not be applied to series like rates or variation in percentages.

$$
\begin{aligned}
\ln y_t = {} & e^{-\lambda\tau} \ln y_{t-\tau} + (1 - e^{-\lambda\tau}) \frac{\alpha}{1-\alpha-\beta} \ln s_k \\
& - (1 - e^{-\lambda\tau}) \frac{\beta}{1-\alpha-\beta} + \ln s_h (1 - e^{-\lambda\tau}) \frac{\alpha}{1-\alpha-\beta} \ln(n+g+\delta) \\
& + (1 - e^{-\lambda\tau}) \ln A + g(t - e^{-\lambda\tau}(t-\tau))
\end{aligned} \tag{1}
$$

where τ refers to the time period and λ is the convergence rate, s_k is the fraction of income invested in physical capital, s_h the fraction invested in human capital, $y = Y/AL$, $k = K/AL$, $h = H/AL$ and δ is the depreciation of the capital. Human and physical capital are subject to decreasing marginal returns and depreciate at the same rate. The level of technology A as well as labour L grow at the exogenously given rates g and n. $0 < \alpha < 1$ and $0 < \beta < 1$ are respectively the share of physical capital and human capital in total output.

This cross-section specification was extended to a panel as in Islam (1995). Letting $\beta = -(1 - e^{-\lambda\tau})$ denote the parameter on income, the speed of convergence is then given by:

$$
\lambda = -\frac{\ln(\beta+1)}{\tau}
$$

3.1 The dataset

Our sample covers 23 regions of Argentina and refers to the 1980-2002 period. All variables are expressed in logarithm. The dependent variable is measured by per capita GDP ($y_{i,t} = GDP/POP$) at constant 1993 thousands of Argentinian pesos and is provided by the CEPAL office of Buenos Aires.

Explanatory variables are measured as an average of the five sub-periods[10]:

- The saving rate s_k at time $t-\tau$ is proxied by the ratio of domestic investment to GDP taken as an average over the five years preceding t (from Provinfo[11]).

- $n_{i,t-\tau}$ is the average population growth rate (from Provinfo) over the period plus 0,05 (see MRW), where 0,05 represents the sum of a common exogenous rate of technical change (g) and a common depreciation rate (δ).

- The investment in human-capital-enhancing activities s_h at time $t-\tau$ corresponds to the number of persons who reached the secondary school over the total population taken as an average over the five years preceding t. This variable is constructed following Audenis *et al.* (2001). Contrary to Caselli *et al.* (1996), Audenis *et al.* (2001) model human capital accumulation in a way that makes it analogous to physical capital one. They take "the amount of total population for whom the decision is to be made at the beginning of every period, whether they should be schooled (i.e. "saved" income to increase productivity next period) or not (i.e. "consumed" income immediately as unskilled workers)".

We also add two control variables in our regressions :

- the export ratio or trade openness variable proxied by the ratio of exports (from Provinfo) to GDP. Because of unavailability of import data, we could not take into account the usual ratio of (Export+Import) to GDP representing a country openness to trade.

- The agglomeration variable is measured by location quotients given by:

$$
Z_i = E_i/R_i
$$

[10]The five considered periods are 1983-1988, 1988-1993, 1993-1998 and 1998-2002. Due to a lack of data on investment and exports, the first period only starts in 1983.

[11]Provinfo is the data collecting department of the Secretaría de Provincias in Buenos Aires.

where E_i and R_i are respectively the employment share[12]) and the reference variable share of province i. Indeed, employment is probably the most directly policy relevant and intuitive measure of the size of an industrial sector. Location quotient are defined following the construction of Gini index as in Krugman (1991):

$$G = \frac{2}{N^2 \bar{Z}} \times \left[\sum_{i=1} \theta_i (Z_i - \bar{Z}) \right] \tag{2}$$

N is the number of provinces. θ_i is the rank of the province i for Z_i. $\bar{Z}$ is the mean of Z_i. This index ranges from 0 (no concentration) to 1 (maximal concentration which means that 100% of the activities is concentrated in a single location).

We retain two reference variables. The first one is the square mile share of province i; Hence,$Z_i AREA$ is a proxy for the provincial share of employment per square mile. The second one refers to the population share of province i; Then, $Z_i POP$ is the provincial relative employment per capita.

3.2 First insights of the convergence analysis

In the MERCOSUR context and in the prospect of the FTAA, the political strength of Argentina is a crucial condition to succeed in the economic integration process of the continent. Hence, the per capita income convergence of Argentina provinces is a fundamental issue for political and social cohesion in the country. It is all the more important than strong regional inequalities are observed. Numerous empirical studies evaluated the per capita income convergence process in Argentina but they were achieved prior to the mid of the eighties. Hence, in this paper, we also update this analysis.

The contribution of regions to Argentinian GDP[13] in 1980, 1990, 1995 and 2002 is illustrated in figure 1. This figure reveals strong regional inequalities and the predominant role of the Buenos Aires region. Indeed, the city of Buenos Aires represents more than 60 % of the total GDP for the considered period followed by Córdoba and Santa Fé that concentrate 10% of the national GDP. Considering the heterogeneity between Argentinean provinces, another crucial question is to check the evolution of this asymmetry.

As a consequence, we proceed to a sigma-convergence analysis by estimating the dispersion of the income per inhabitant, i.e. the coefficient of variation of per capita GDP, from 1980 to 2002. In order to control for spatial autocorrelation effects, trends in coefficient of variation on filtered data are also presented[14] following the filtering method proposed in the preceding section. The two trends in coefficient of variation are similar even if it is slightly smoothen at beginning of the nineties when data are filtered.

Trends in figure 2 stress on two distinct periods. First, we distinguish a period of strong divergence between provinces from 1980 to 1988 corresponding the hyper inflation period in Argentina also called the *Lost Decade* in Latin America. Second, a long period of convergence in per capita GDP in Argentina is observed with a sudden decrease in the coefficient of variation in 1990 followed by a gradual decline until 2002.

[12] Only aggregated employment have been annually collected overall the period.

[13] NOA region (Nord Ouest Argentin) covers Catamarca, Jujuy, Santiago del Estero and Tucumán provinces, NEA (Nord Est Argentin) region covers Formosa and Chaco, Argentinian Mesopotamia in the North coast covers Entre Ríos, Corrientes and Misiones, Nuevo Cuyo region close to the Cordillera de los Andes covers Mendoza, San Juan and La Rioja, the Easy Center counts for Córdoba and Santa Fé, the Center covers Buenos Aires province and La Pampa, finally Patagonia is located in the South of Argentina with Neuquén, Río Negro, Santa Cruz and Tierra del Fuego.

[14] With $\gamma_E = 0,4$.

Figure 1: Regional Share in Total Argentinian GDP, 1980, 1990, 1995 and 2002

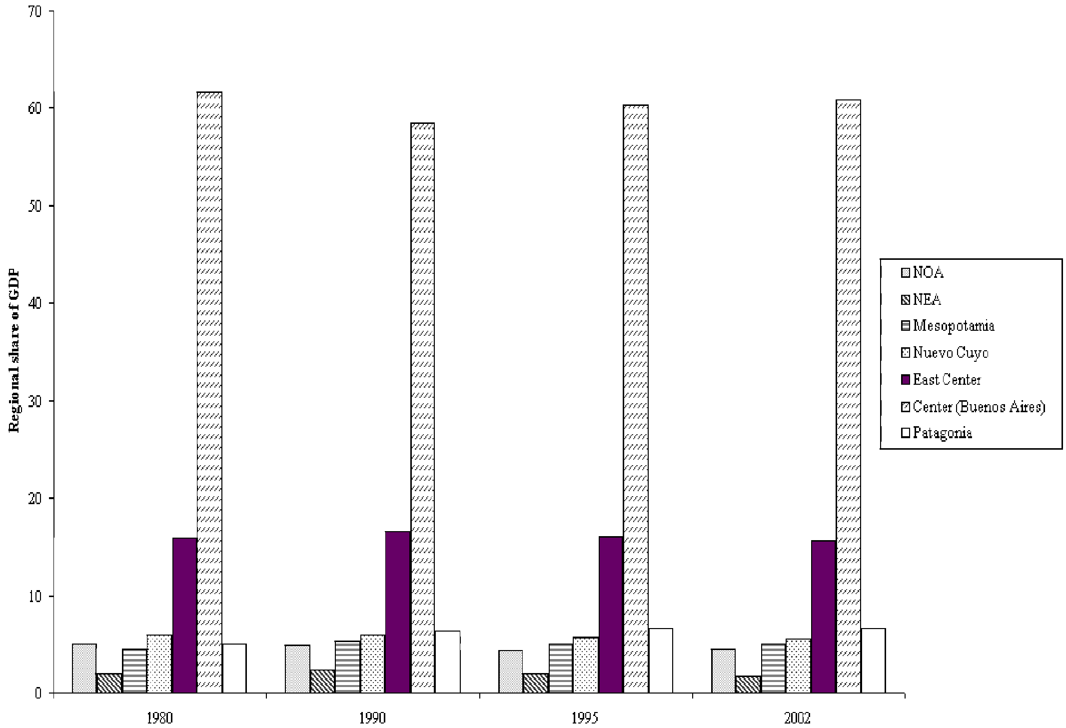
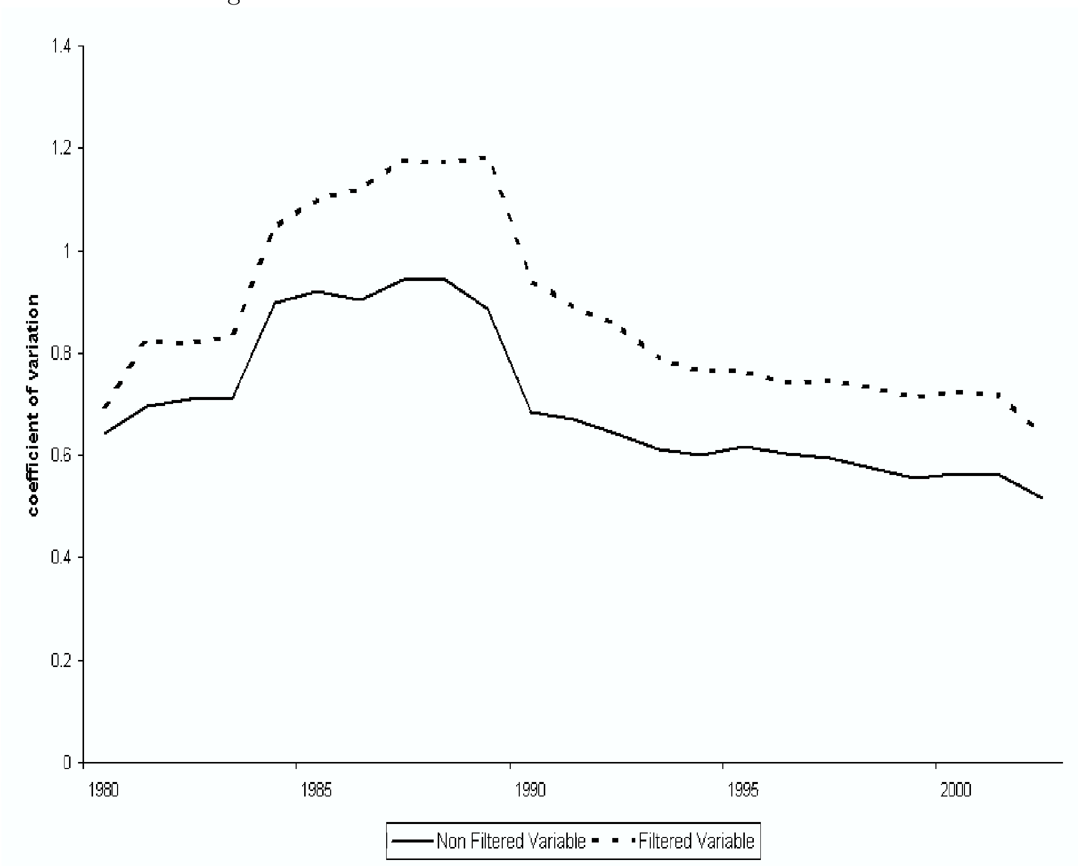


Figure 2: Coefficient of variation between 1980 and 2002



3.3 A dynamic panel data approach for conditional β convergence test

Convergence studies were originally based on cross-sections and estimated using OLS (Barro and Sala-i-Martin, 1995). Later on, the framework of cross-sections studies was very criticized. Indeed, the initial level of technology, that should be included in a conditional convergence specification, is not observed. Since it is also correlated with another regressor (initial income), all cross-sections studies suffer from an omitted variable. Thus, Islam (1995) proposed to set up a convergence analysis in a panel data framework (within-group estimator) where it is possible to control for individual-specific time invariant characteristics of countries (like the initial level of technology) using fixed effects. However, whether the potential advantage can be realized largely depends on the panel data estimators used, and in the case of an endogeneity correction, the availability of feasible instruments.

3.3.1 The First-differenced GMM estimator

From different procedures suggested in the literature for dynamic panel data models, most studies (Caselli *et al.* (1996), Henderson (2000)) employ the generalized method of moments estimator in first differences (GMM-DIFF) suggested by Arellano and Bond (1991). In growth analysis, the GMM estimator was first applied in the paper of Caselli *et al.* (1996). For simplicity, we consider an growth model with unobserved individual-specific effects:

$$\ln y_{i,t} = \beta \ln y_{i,t-1} + \delta \ln X_{i,t} + \eta_i + \epsilon_t + \upsilon_{i,t} \text{ where } |\beta| < 1 \tag{3}$$

where $y_{i,t}$ is the per capita income of region i at the date t, $X_{i,t}$ is a vector of economic growth determinants, $i = 1, ..., N$ and $t = 1, ..., T$, η_i is the individual specific effect, ϵ_t is a time constant and $\upsilon_{i,t}$ the standard error. The time index t refers to an interval of five years.

Following hypothesis are respected: $E[\eta_i] = 0$ and $E[\upsilon_{i,t}] = 0$, $E[\upsilon_{i,t}\eta_i] = 0$ for $i = 1, ..., N$ and $t = 2, ..., T$. The constant term ϵ_t is particularly considered in growth equations:

> "the inclusion of time dummies allows for common long-run growth in per capita GDP consistent with common technical progress without violating the validity of the additional moment restriction used by the system GMM estimator" (Bond *et al.*, 2001).

We transform all the variable as deviation from time means[15] following Caselli *et al.* (1996). This eliminates the need for time dummies[16].

The first step in the estimation procedure is to eliminate the individual effects via a first-difference transformation:

$$\ln y_{i,t} - \ln y_{i,t-1} = \tilde{\beta}(\ln y_{i,t-1} - \ln y_{i,t-2}) + \tilde{\delta}(\ln X_{i,t} - \ln X_{i,t-1}) + (\upsilon_t - \upsilon_{t-1}) \tag{4}$$

As instruments for the lagged difference of the endogenous variable - or other variables which are correlated with the differenced error term, all lagged levels of the variable in question are used, starting with lag two and potentially going back to the beginning of the sample. Consistency of the GMM estimator requires a lack of second order serial correlation in the residuals of the differenced specification. The overall validity of instruments can be checked with a Sargan test of over-identifying restrictions (see Arellano and Bond, 1991).

We can also write equation (4) with the following form:

$$\Delta \ln y_{i,t} = \tilde{\beta}\Delta \ln y_{i,t-1} + \tilde{\delta}\Delta \ln X_{i,t} + \Delta\upsilon_{i,t} \text{ for } t = 3, ..., T \text{ and } i = 1, ..., N \tag{5}$$

where $y_{i,t-2}$ and all previous lags are used as instruments for $\Delta \ln y_{i,t-1}$ assuming that $E[\upsilon_{it}\upsilon_{is}] = 0$ for $i = 1, ..., N$ and $s \neq t$ and that initial conditions on $\ln y_{i1}$ are predetermined as $E[\ln y_{i1}\upsilon_{it}] = 0$ for $i = 1, ..., N$ and $t = 2, ..., T$.

[15] Corresponding to the mean of the N Argentinian provinces at each period.

[16] Corresponding to the time constant ε_t in the growth equation (3).

Together, these assumptions imply the following $m = 0.5(T-1)(T-2)$ moment restrictions: $E[Z_i' \Delta \upsilon_i] = 0$. Z_i is the $(T-2) \times m$ matrix given by :

$$Z_i = \begin{bmatrix} \ln y_{i,1} & 0 & 0 & ... & 0 & ... & 0 & \ln X_{i,1} \\ 0 & \ln y_{i,1} & \ln y_{i,2} & ... & 0 & ... & 0 & \ln X_{i,2} \\ ... & . & ... & ... & . & ... & . & . \\ 0 & 0 & 0 & ... & \ln y_{i,1} & ... & \ln y_{i,T-2} & \ln X_{i,T-2} \end{bmatrix}$$

where $X_{i,T-k} = (\ln(I/GDP)_{i,1}, ..., \ln(I/GDP)_{i,t-k}, \ln(n+g+\delta)_{i,1}, ..., \ln(n+g+\delta)_{i,t-k},$ $\ln(SCHOOL2)_{i,1}, ..., \ln(SCHOOL2)_{i,t-k}, 1)$. $\Delta \upsilon_i$ is the vector $(\Delta \upsilon_{i,3}, \Delta \upsilon_{i,4},\ \Delta \upsilon_{i,4}, ..., \Delta \upsilon_{i,T})'$ of $(T-2)$ dimension. This yields a consistent estimator of $\tilde{\beta}$ as $N \to \infty$ with T fixed.

However, this first-differenced GMM estimator has been found to have poor finite sample properties, in terms of bias and imprecision, in one important case. This occurs when the lagged levels of the series are only weakly correlated with subsequent first-differences, so that the instruments available for the first differenced equations are weak.

The GMM estimator in first differences has been criticized recently given that annual log of per capita GDP is likely to be persistent[17]. Weak correlation exists between the growth rate of log per capita GDP and the lagged log of per capita GDP levels. Lagged levels of the series provide weak instruments for first differences in this case.

3.3.2 The System GMM Estimator

More recent studies (Bond *et al.* (2001), Badinger *et al.* (2004), Yudong and Weeks (2003)) obtain new results from dynamic panel data econometrics by using system GMM (GMM-SYS) estimators as proposed by Arellano and Bover (1995) and Blundell and Bond (1998) to overcome the problem of weak instruments observed in GMM-DIFF[18].

Blundell and Bond (1998) show that biases can be dramatically reduced by introducing lagged first-difference as instruments for equations in levels, in addition to the usual lagged levels as instruments for equations in first-differences. They propose a system GMM estimator, where a system of equations is estimated in first differences and in levels. The $(T-2)$ differences equations, given by (5) are supplemented by the following $(T-1)$ levels equations:

$$\ln y_{i,t} = \beta \ln y_{i,t-1} + \delta \ln X_{i,t} + \eta_i + \upsilon_{i,t} \text{ with } t = 2, ..., T \text{ and } i = 2, ..., N \tag{6}$$

where lagged first differences are used as instruments for additional equations, based on two new assumptions. (1) $E[\mu_i \Delta \ln y_{i,2}] = 0$ for $i = 1, ..., N$, stating that the dependent variable in first difference with $t = 2$ is not correlated with the individual effect; (2) $E[\mu_{it} \Delta \ln y_{i,t-1}] = 0$ and $E[\mu_{it} \Delta \ln X_{i,t}] = 0$ for $i = 1, ..., N$, $t = 3, 4, ..., T$ and $\mu_{i,t} = \mu_i + \upsilon_{i,t}$ indicating that first difference regressors are not correlated with the error term.

The condition (2) allows to use first differences of the series as instruments for equations in levels (Arellano et Bover, 1995). The instrument matrix for equations in levels can then be written as:

$$Z_i^+ = \begin{bmatrix} \Delta \ln y_{i,1} & 0 & 0 & ... & 0 & ... & 0 & \Delta \ln X_{i,1} \\ 0 & \Delta \ln y_{i,1} & \Delta \ln y_{i,2} & ... & 0 & ... & 0 & \Delta \ln X_{i,2} \\ ... & . & ... & ... & . & ... & . & . \\ 0 & 0 & 0 & ... & \Delta \ln y_{i,1} & ... & \Delta \ln y_{i,T-2} & \Delta \ln X_{i,T-2} \end{bmatrix}$$

[17] This bias problem has given very important rate of convergence (3% to almost 10% in Caselli *et al.* (1996)).

[18] These studies find more reasonable rates of convergence ranged from 2% to 4%.

where $X_{i,T-k} = (\Delta\ln(I/GDP)_{i,1}, ..., \Delta\ln(I/GDP)_{i,t-k}, \Delta\ln(n+g+\delta)_{i,1}, ..., \Delta\ln(n+g+\delta)_{i,t-k},$ $\Delta\ln(SCHOOL2)_{i,1}, ..., \Delta\ln(SCHOOL2)_{i,t-k}, 1)$.

Hence, the complete instrument matrix for the GMM-SYS estimator is:

$$Z_i^S = \begin{bmatrix} Z_i & 0 \\ 0 & Z_i^+ \end{bmatrix}$$

where Z_i is given by matrix (3.3.1) and Z_i^S by matrix 3.3.2[19]. Instead of using robust variances from the first step for the second step of GMM-DIFF and GMM-SYS, a correction of the second step robust variance based on Weidmeijer (2000) is used.

3.4 Results of estimation

We can rewrite the testable equation (1) as follows:

$$\ln y_{it} = \gamma \ln y_{i,t-1} + \beta_1 \ln sk_{it} + \beta_2 \ln sh_{it} + \beta_3 \ln(n+g+\delta)_{it} + \eta_i + \epsilon_t + \upsilon_{it} \tag{7}$$

$I = 1, ..., N$ and $t = 2, ..., T$. It represents a general panel dynamic framework with $\gamma = e^{-\lambda\tau}, \beta_1 = \beta = (1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}, \beta_2 = -\beta$. $\eta_i = (1-e^{-\lambda\tau})\ln A$ corresponds to the region specific effect, $\epsilon_t = g(t - e^{-\lambda\tau}(t-\tau))$ denotes the time specific effect and υ_{it} is the error term usually assumed $IID(0, \sigma^2), \tau = 5$ years (the period assumed to let the region to converge to its steady state).

To evaluate the performance of the GMM-SYS estimator, we establish the same bound for the autoregressive parameter as in Caselli *et al.* (1996) or Bond *et al.* (2001) in the growth models analysis. The bias observed in the pooled OLS levels and the within groups estimator in models with fixed T is used as a reference to define an approximate upper and lower bound for the autoregressive parameter in the growth regressions. As Hsiao (1986) shows, omitting unobserved time invariant country effects in a dynamic panel data model will cause OLS levels estimate to be biased upward and inconsistent, given the positive correlation between the lagged dependent variable and the permanent country fixed-effect. At the opposite, an alternative estimation technique which takes into account the unobserved country specific effects is the within group estimator. The within groups estimator produces a downward bias with the extent of attenuation increasing when exogenous covariates are added (Nickell, 1981).

The unobserved country-specific effects (μ_i) reflect differences in the initial level of efficiency, whilst the period-specific intercepts (η_t) capture productivity changes that are common to all countries as in MRW. In equation (6), there is no period-specific intercepts as required for the estimation of an empirical growth model. Thus, as pointed by Bond *et al.* (2001), "the inclusion time dummies allows for common long-run growth in per capita GDP consistent with common technical progress without violating the validity of the additional moment restriction used by the system GMM estimator". As in their paper and in Caselli *et al.* (1996), we change all variable into deviations from time means which correspond to means across the N individual Argentinian provinces for each period. By this way, we do not need to include time dummies in GMM regressions.

In tables 2 on unfiltered data and 3 on filtered data, results from a GMM-DIFF and a GMM-SYS are compared to a within groups (WG) and a pooled OLS estimation as in Bond *et al.* (2001) and Yudong and Weeks (2000). Indeed, performance of the GMM-SYS and the GMM-DIFF can be tested by the identification of an estimation range for the convergence speed provided by the OLS and the within groups estimators. All the explanatory set is considered to be endogenous and thus instrumented.

[19] The complete set of second-order moment conditions available given assumption (1) and can be expressed as $E[Z_i^+{}' u_i^+] = 0$ where $u_i^+ = ((\Delta\upsilon_{i3}, \Delta\upsilon_{iT}, \upsilon_{i3}, ..., \upsilon_{iT})'$.

Table 2: Estimations of simple and augmented Solow Models from 1983 to 2002-unfiltered data

Dependent variable is ln (y_{it})								
	OLS		WG		GMM-DIF		GMM-SYS	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
$\ln(y_{i,t-1})$	0.986***	0.979***	0.143	0.184*	0.591	0.302	1.023***	1.012***
	(0.006)	(0.035)	(0.09)	(0.098)	(0.574)	(0.45)	(0.041)	(0.133)
$\ln(I/GDP)_{i,t}$	0.119***	0.119***	0.029	0.032	0.142*	0.174	0.223***	0.145*
	(0.041)	(0.041)	(0.06)	(0.06)	(0.081)	(0.105)	(0.064)	(0.08)
$\ln\,(n_{i,t}+g+\delta)$	-0.08	-0.074	-0.114*	-0.123**	-0.072	0.093	-0.183**	-0.162**
	(0.061)	(0.069)	(0.058)	(0.057)	(0.196)	(0.089)	(0.079)	(0.076)
$\ln(SCHOOL2)_{i,t}$		0.019		0.294		1.478		0.236
		(0.094)		(0.276)		(0.955)		(0.947)
Implied λ	0.282%	0.424%		33.856%				
Observations	92	92	92	92	69	69	92	92
R-squared	0.88	0.88	0.02	0.07				
Hansen test					0.604	0.511	0.613	0.295
m1 test					0.575	0.909	0.027	0.1
m2 test					0.16	0.903	0.151	0.213

Robust standard errors in parentheses and two-step estimator

Figures reported for the tests are p-values. λ is the annual rate of convergence.

* significant at 10%; ** significant at 5%; *** significant at 1%

Instruments used for GMM-DIFF are $\ln(y_{i,t-2})$,$\ln(I/GDP)_{i,t-1}$,ln $(n_{i,t-1}+g+\delta)$ and $ln(SCHOOL2)_{i,t-1}$.

Instruments used for SYS-DIFF are $\Delta ln(y_{i,t-2})$,$\Delta ln(I/GDP)_{i,t-1}$, $\Delta ln(n_{i,t-1}+g+\delta)$ and $\Delta ln(SCHOOL2)_{i,t-1}$.

Table 3: Estimations of simple and augmented Solow Models from 1983 to 2002-filtered data

Dependent variable is ln (y_{it})								
	OLS		WG		GMM-DIF		GMM-SYS	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
$\ln(y_{i,t-1})$	0.975***	0.966***	0.102	0.132*	1.593**	1.426	0.971***	0.961***
	(0.006)	(0.046)	(0.112)	(0.123)	(0.767)	(0.927)	(0.275)	(0.251)
$\ln(I/GDP)_{i,t}$	0.163***	0.165***	0.058	0.002	0.275	0.184	0.184*	0.093
	(0.036)	(0.037)	(0.061)	(0.067)	(0.286)	(0.401)	(0.089)	(0.069)
$\ln\ (n_{i,t} + g + \delta)$	-0.022	-0.017	-0.011	-0.061	-0.796	-0.94	-0.238**	-0.148*
	(0.044)	(0.046)	(0.082)	(0.085)	(0.856)	(0.818)	(0.092)	(0.085)
$\ln(SCHOOL2)_{i,t}$		0.036		-0.240*		0.294		0.065
		(0.127)		(0.13)		(2.774)		(0.489)
Implied λ	0.506%	0.692%		40.499%			0.589%	0.796%
Observations	91	91	91	91	68	68	91	91
R-squared	0.89	0.89	0.02	0.07				
Hansen test					0.224	0.206	0.217	0.125
m1 test					0.175	0.256	0.079	0.098
m2 test					0.411	0.713	0.205	0.282

Robust standard errors in parentheses and two-step estimator

Figures reported for the tests are p-values

* significant at 10%; ** significant at 5%; *** significant at 1%

Instruments used for GMM-DIFF are $\ln(y_{i,t-2})$,$\ln(I/GDP)_{i,t-1}$,ln $(n_{i,t-1} + g + \delta)$ and $ln(SCHOOL2)_{i,t-1}$.

Instruments used for SYS-DIFF are $\Delta ln(y_{i,t-2})$,$\Delta ln(I/GDP)_{i,t-1}$, $\Delta ln(n_{i,t-1} + g + \delta)$ and $\Delta ln(SCHOOL2)_{i,t-1}$.

The selection of the spatial filtering parameter γ_E is based on a variation of the distance decay parameter (weight matrix). The fit and tests of the model are used to identify appropriate spatial weights. Thus, the chosen model, i.e. the selected distance decay, provides the best fit simultaneously capturing, if possible, the overall spatial interaction that characterizes the change in regional per capita income[20]. Table 3 allows us to compare filtered to non filtered data.

The table also gives three tests of the appropriateness of instruments for GMM-DIFF en GMM-SYS estimators. The Hansen test is a test of over-identifying restrictions under the null hypothesis of the validity of instruments (Arellano and Bond, 1991, 1998). The high p-values suggest in both estimators that we cannot reject the null hypothesis and the set of instruments is then appropriate in all models. The two latter are tests of first order and second-order serial correlation in the first differenced residuals. Reported p-values give the probability of correctly rejecting the null hypothesis of no autocorrelation. As required, the test in the GMM-SYS for first order autocorrelation rejects the null hypothesis while tests for second-order autocorrelation fails to reject the null hypothesis of no autocorrelation. However, column five in both tables 2 and 3 do not reject the null hypothesis of no autocorrelation meaning that GMM-DIFF is not correctly instrumented. Indeed, first-differencing introduces MA(1) serial correlation when time-varying component of the error term in levels is serially uncorrelated. Therefore GMM-SYS estimator is consistent only when second-order correlation is not significant although first-order correlation need not be zero.

By correcting dynamic panel bias, GMM-SYS and GMM-DIFF should better perform estimations of the autoregressive variable y_i for empirical growth models. Nevertheless, as in Bond *et al.* (2001), our tests (Hansen test and tests for first and second order autocorrelation) only validate the GMM-SYS for both models. The validity of the new instruments is not rejected (i.e. instruments for equations in levels are lagged first-differences of variables). Comparing column (6) and (8), coefficients yielded by GMM-DIFF and GMM-SYS estimators remain stable.

It is worth noting that coefficients on initial per capita GDP are very close to unity. Thus, we wonder if series on initial per capita GDP are not highly persistent. We then compute a unit root test on GDP per capita series that reject the presence of a unit root. Blundell and Bond (2000) compared the performance of GMM-DIFF and GMM-SYS in presence of an autoregressive coefficient around 0,9 and concluded that GMM-SYS yielded much more reasonable parameter estimates particularly in presence of autoregressive coefficient very close to unity.

We first look at the results when spatial dependence is not considered. A positive coefficient less than unity on initial per capita GDP is interpreted as conditional convergence. Columns [1], [2] and [3], [4] respectively report OLS and WG estimators results. Columns [5], [6] and [7], [8] present estimated parameters using respectively the GMM-DIFF and the GMM-SYS estimators. We notice that OLS coefficients are slightly lower than unity indicating a limited convergence process between Argentinean provinces. The coefficient goes down to 0,18 with the WG estimator. Hence, it varies in the expected way. The GMM-DIFF coefficient falls between the upper and the lower bound but appears to suffer from weak instruments problem according to *m1* tests that reject the presence of MA(1) autocorrelation (see columns [5] and [6]). The GMM-SYS estimator also yields a coefficient that exceeds the OLS one which may be due to the weak performance of instruments (see the *m1* test in column [8]) and to the presence of the spatial autocorrelation between variables. It appears that those results are potentially misleading due to a model specification and that we have to take spatial dependencies into account.

Comparing coefficient values obtained with filtered variables, the consideration of spatial dependence obviously has a substantial impact on convergence. Besides, it is worth noting that coefficients of GMM-SYS estimator falls within the upper and the lower bound given by OLS and WG estimators. Surprisingly, the GMM-DIFF estimator yields upwards biased coefficients and tests on instruments unvalidate them. The GMM-SYS estimator

[20]Here $\gamma_E = 0,4$.

conclude to a divergence process when data are not filtered and this process turns to be significantly convergent when spatial autocorrelation is removed. When we control for spatial autocorrelation between provinces (i.e., when we remove the influence of geographic proximity between provinces), convergence test is validated by the GMM-SYS estimator. The influence of Buenos Aires province on neighbouring provinces appears to upward biases the autoregressive coefficient on unfiltered variables in columns [7] and [8] of table 2.

Our results concluding to a conditional convergence process within Argentinian provinces are in line withe those of Utrera and Koroch (1998) and Marina (2001) who applied OLS estimates to the Argentinian case but for previous periods. However, rates of convergence are extremely slow as they are below 1% (0,59% in the simple Solow model and 0,80% in the augmented model) and lower than those obtained by Bond *et al.* (2001), Yudong and Weeks (2003) and Badinger *et al.* (2004) from 2% to 6%. This may probably explained by our panel dataset since growth in Argentina provinces was undoubtedly more unstable than growth in European Union and Chinese regions during the 1983-2002 period. Indeed, except during the 1991-1997 period characterized by macroeconomic stability and by important per capita GDP growth rates (reaching 4,69% in 1994), these latter remained very unstable and often negative over the whole period.

Concerning other variables, we notice that as expected by the simple Solow model, the investment rate impacts positively and significantly the growth rate mainly by promoting industrial development of a region. Nevertheless, the addition of investment on secondary school variable makes the investment rate insignificant in the augmented Solow model. A possible explanation may be that Argentinian education policies were not set as much a priority as in East Asian countries during the last two decades.

4 Regional convergence and Geography

4.1 An overview of economic geography and convergence studies

New economic geography theories aim at explaining the location choices of firms and their agglomeration process (Krugman, 1991). The geographical distribution of activities is rarely random: it can be explained by natural conditions, history as well as by the presence of other firms already located in a specific region. Then, opposite forces of concentration (the so-called backward and forward linkages) drive to a cumulative and endogenous process of agglomeration (see Fujita *et al.*, 1999).

The effect of uneven spatial distribution of activities on regional economic growth has also been pointed out in new economic geography (Baldwin *et al.* (2003), Martin and Ottaviano (1999, 2001)). This theory is based on several determinants affecting the characteristics of the production process (such as increasing returns to scale, monopolistic competition, vertical linkages[21], externalities, technological spillovers[22]...). Hence, agglomeration is considered as a factor of growth. As sum up by Baumont *et al.* (2000), economic geography theory has evidenced that concentration of activities favours growth and, at the same time, that economic growth may be considered as a new agglomeration force. This theoretical approach heeds the relation between convergence process among regional economies and agglomeration.

Nevertheless, few empirical studies on regional economic convergence applying Barro growth equations consider the effect of explanatory concentration variables. In Madariaga Montout and Ollivaud (2004), an empirical analysis on the MERCOUSR and the NAFTA investigate and sheds light on the existence of a positive relation between the growth rate and concentration of activities. Henderson (2000) explores if urban over concentration really

[21] See Venables (1996) for a further definition of these linkages.

[22] In Martin and Ottaviano (1999), local and global spillovers are distinguished in production processes to determine their impacts on growth.

affects economic growth rates in a robust and consistent fashion using data from a panel of 80 to 100 countries every 5 years from 1960 to 1995. He finds that not only is there an optimal degree of urban concentration that varies with country income, but departures from optimal concentration result in substantial growth losses. Hence, urban concentration increases sharply as income rises, up to a critical value of the per capita income[23], before declining modestly. Countries with highly excessive concentration include Argentina. As a matter of fact, we can easily observe a strong spatial order in Argentinian regions since richest regions (Buenos Aires, Santa Fé and Córdoba provinces) are also the most geographically concentrated.

Besides, a spate of empirical cross-country studies by Slaughter (1999), Edwards (1998) and Coe *et al.* (1997) suggested that the impact of trade liberalization in goods on the long run rate of economic growth is positive, although a more recent paper (Rodriguez and Rodrik (1999)) questioned the robustness of these results. Hence, we also control for the impact of openness[24] in the estimation by introducing an Exports/GDP variable in the explanatory set.

Tables 4 and 5 report regressions on unfiltered and filtered variables taking into account the incidences of trade liberalization (proxied by an export ratio) and agglomeration. We only report results of the GMM-SYS estimator as it proved to be the best estimator in the previous section.

4.2 Results of the estimation

The coefficient of the autoregressive variable $y_{i,0}$ confirms a conditional convergence process but coefficients on filtered data are substantially lower than on unfiltered data (except in column [4] in table 5)[25]. Convergence rates obtained in table 5 are closer to 4% as in Bond *et al.* (2001), Yudong and Weeks (2003) and Badinger *et al.* (2004). As a matter of fact, this increase in the rate of convergence is partly explained by the addition of agglomeration variables. It indicates the importance of inequalities in the distribution of activities in Argentina. Indeed, when we control the impact of concentration, the results of convergence are similar to other empirical works. Therefore,agglomeration of activities seem to be an important structural characteristic explaining differences in steady states.

Another important contribution of the withdrawal of spatial correlation is that the agglomeration of relative employment per capita, Z_iPOP, becomes significant in column [1] and [3] as expected in the empirical literature. Then, removing spatial autocorrelation between provinces partly controls for the effect of proximity to Buenos Aires province. This allows us to really estimate pure agglomeration impact on growth but not geographical proximity dependencies. Coefficients of Z_iAREA are never significant showing the lesser relevance of referring the agglomeration variable to employment per square mile than to employment per capita. Z_iPOP estimates the activity agglomeration relatively to the market size in terms of consumers. It emphasizes the great impact of the labour force proximity for economic growth in all Argentinean provinces and therefore for the per capita income convergence.

The export variable does not influence significantly the growth of per capita income even if the sign of the coefficient is positive. One possible explanation is the measure of this openness variable since we did not include imports in the variable as usually done in the literature. Thus, imports data could provide a complete information about the different stages of liberalization in Argentina. From 1983 to 2002, we distinguish two important steps

[23] In average equals to 5 000 US$.

[24] Another alternative is to follow Caselli *et al.* (1996) who introduced a rate of change in the terms of trade that captures favorable shocks to external competitiveness.

[25] Once again, we are in presence of an autoregressive coefficient around 0,9 in tables 4 and the presence of a unit root is also rejected by the tests.

Table 4: Convergence and geography estimations from 1983 to 2002- unfiltered

Dependent variable is ln (y_{it})				
	GMM-SYS			
	[1]	[2]	[3]	[4]
$\ln(y_{i,t-1})$	0.984***	0.990***	0.973***	0.991***
	(0.064)	(0.058)	(0.082)	(0.063)
$\ln(I/GDP)_{i,t}$	0.242***	0.194***	0.243***	0.171***
	(0.057)	(0.065)	(0.052)	(0.05)
ln $(n_{i,t} + g + \delta)$	-0.139	-0.249**	-0.121	-0.242**
	(0.21)		(0.249)	
$ln(Z_iPOP)_{i,t}$	0.122		0.098	
	(0.21)		(0.249)	
$ln(Z_iAREA)_{i,t}$		-0.097		-0.069
		(0.063)		(0.051)
ln $(EXP/GDP)_{i,t}$			0.018	0.003
			(0.021)	(0.02)
Implied λ	0.323%	0.201%	0.547%	0.181 %
Observations	87	87	86	86
Hansen test	0.814	0.857	0.937	0.901
m1 test	0.005	0.014	0.008	0.022
m2 test	0.284	0.285	0.316	0.339

Robust standard errors in parentheses and two-step estimator
Figures reported for the tests are p-values
* significant at 10%; ** significant at 5%; *** significant at 1%
Instruments used for GMM-DIFF are $\ln(y_{i,t-2})$,$\ln(I/GDP)_{i,t-1}$, ln $(n_{i,t-1} + g + \delta)$ and the lag of first difference of added variables.
Instruments used for SYS-DIFF are $\Delta ln(y_{i,t-2})$,$\Delta ln(I/GDP)_{i,t-1}$, $\Delta ln(n_{i,t-1} + g + \delta)$ and the lagged level of added variables.

Table 5: Convergence and geography estimations from 1983 to 2002- filtered

Dependent variable is ln (y_{it})				
	GMM-SYS			
	[1]	[2]	[3]	[4]
$\ln(y_{i,t-1})$	0.826***	0.742***	0.817***	1.048**
	(0.177)	(0.19)	(0.185)	(0.456)
$\ln(I/GDP)_{i,t}$	0.131*	0.115*	0.135*	0.041
	(0.07)	(0.058)	(0.066)	(0.077)
ln $(n_{i,t} + g + \delta)$	-0.117*	-0.079	-0.108	-0.102
	(0.062)	(0.049)	(0.072)	(0.125)
$ln(Z_iPOP)_{i,t}$	0.282*		0.316*	
	(0.159)		(0.159)	
$ln(Z_iAREA)_{i,t}$		0.07		-0.053
		0.07		-0.053
ln $(EXP/GDP)_{i,t}$			0.006	0.029
			(0.031)	(0.045)
Implied λ	3.823%	5.968%	4.042%	
Observations	87	87	86	86
Hansen test	0.115	0.1	0.371	0.079
m1 test	0.03	0.037	0.027	0.197
m2 test	0.335	0.296	0.365	0.949

Robust standard errors in parentheses and two-step estimator
Figures reported for the tests are p-values
* significant at 10%; ** significant at 5%; *** significant at 1%
Instruments used for GMM-DIFF are $\ln(y_{i,t-2})$,$\ln(I/GDP)_{i,t-1}$, ln $(n_{i,t-1} + g + \delta)$ and the lag of first difference of added variables.
Instruments used for SYS-DIFF are $\Delta ln(y_{i,t-2})$,$\Delta ln(I/GDP)_{i,t-1}$, $\Delta ln(n_{i,t-1} + g + \delta)$ and the lagged level of added variables.

in Argentina trade openness, the first one corresponding to the free trade area signed with Brazil in 1985 and the second one to the signature of the MERCOSUR in 1991.

5 Conclusion

This paper extends the analysis of previous works on Argentina's provincial income convergence by computing system and first-differenced estimators from 1983 to 2002. The literature on income convergence in Argentina only used the traditional approach of Barro and Sala-i-Martin (1995) or the Within-Group model estimator to deal with a panel dataset. The aim of this paper is to apply a new method of estimating convergence that combines spatial filtering and panel data econometrics. As a consequence, we adopt the methodology used by Badinger *et al.* (2004) using a dynamic panel data. We employ a two-step procedure. The first step applies a filtering technique as proposed by Getis and Griffith (2002) to remove spatial autocorrelation from variables. Variables were filtered before the computation of regressions. The second one estimates the speed of convergence with two different GMM estimators. We use more informative sets of instruments than in Caselli *et al.* (1996) by using the system GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998).

We confirm the important result obtained by Bond et al. (2001) that using system GMM estimator is preferable to the first-differenced estimator in empirical growth works. We also introduced other variables as instruments that are not included in the augmented Solow model, for example the lags of agglomeration variables. When testing for the simple and augmented Solow models, both GMM-DIFF and GMM-SYS estimators conclude to a conditional convergence when data are filtered. Our results highlight the importance of removing spatial autocorrelation, as we find a conditional divergence when data are unfiltered. We also confirm that investment rate impacts positively the growth rate of per capita GDP as it should promote the development of a region. However, investment in human capital has no significant influence on the growth rate. Finally, we add agglomeration and openness variables to control for their impact on growth rate. The coefficient of the autoregressive variable y_i reveals the presence of a conditional convergence process. Only coefficients on agglomeration variables are positive and significant. This result emphasizes the great impact of the labour force proximity for the economic growth and therefore for the per capita income convergence. Indeed, the localization of economic activities is affected notably by the proximity of the labour market and the consumers.

References

[1] **Abreu M., De Groot H.L.F., Florax R.J.G.M** (2004) "Space and Growth: A Survey of Empirical Evidence and Methods", *Tinbergen Institute Discussion Paper*,TI 2004-129/3.

[2] **Arellano M., Bond S.**, (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*,58, p.277-297.

[3] **Arellano M., Bover S.**, (1995), "Another Look at Instrumental Variable Estimation of Error-component Models", *Journal of Econometrics*, 68, p.29-52.

[4] **Audenis C., Biscoup P., Fourcade N. and Loisel O.,**(2001), "Testing the Augmented Solow Growth Model: An Empirical Reassessment using Panel Data", *Direction des Etudes et Synthèses Statistiques Working paper*, n° G 2001/05, INSEE. 68, p.29-52.

[5] **Badinger, H., W. Mller and G. Tondl** (2004) “Regional Convergence in the European Union (1985-1999): A Spatial Dynamic Panel Analysis”, *Regional Studies*, 38, pp. 241-53.

[6] **Baldwin R., Forslid R., Martin P., Ottaviano G.M. and Robert-Nicoud F.**, (2003), *Global and Local Spillovers Models*, In Economic Geography and Public Policy, Princeton University Press.

[7] **Barro R.J., Sala-i-Martin X.,**(1991), “Convergence Across States and Regions”, *Brooking Papers on Economic Activity,* 1, pp. 107-158.

[8] **Barro R.J., Sala-i-Martin X.**, (1995), *Economic Growth*, Mc GrawHill.

[9] **Baumont C., Ertur C. and Le Gallo J.**, (2000), ”Geographic Spillover and Growth: A Spatial Econometric Analysis for European Regions”, *LATEC - Document de travail*, Economie 2000-07.

[10] **Blundell R. and Bond S.**, (2000), ”GMM Estimation with Persistent Panel Data: An Application to Production Functions”, *Econometric Reviews*, 19(3), p.321-340.

[11] **Blundell R. and Bond S.**, (1998), ”Initial Conditions and Moment Restrictions in Dynamic Panel Data Models”, *Journal of Econometrics*, 87(1), p.115-143.

[12] **Bond S., Hoeffler A. and Temple J.**, (2001), ”GMM estimation of empirical growth models”, *CEPR discussion paper*, n°3048.

[13] **Caselli F., Esquivel G. and Lefort F.**, (1996), ”Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics”, *Journal of Economic Growth*, 1, p.363-389.

[14] **Coe D., Helpman E. and Hoffmaister E.**, (1997), ”North-South R&D Spillovers”, *The Economic Journal*,107, p.134-149.

[15] **Edwards S.**, (1998), ”Openness, Productivity and Growth: What Do We Really Know?”, *Economic Journal*, 108, p.383-398.

[16] **Easterly W., Loaysa N. and Montiel P.**, (1997), ”Has Latin AMerica’s Post-Reform Growth Been Disappointing”, *Journal of International Economics*, 43(3-4), p. 287-311.

[17] **Fingleton B.**, (1999), ”Estimates of Time to Economic Convergence: An Analysis of Regions of the European Union”, *International Regional Science Review*, 22(1), p.5-34.

[18] **Fujita M., Krugman P., Venables A., (1999)**, *The Spatial Economy: Cities, Regions and International Trade*, London School of Economics.

[19] **Getis A. and Griffith D.A.**, (2002), ”Comparative Spatial Filetring in regression Analysis”, *Geographical Analysis*, 24(3), p.189-206.

[20] **Getis A. and Ord J.K.**, (1992), ”The Analysis of Spatial Association by Use of Distance Statistics”, *Geographical Analysis*, 34(2), p.130-140.

[21] **Henderson V.,**(2000), ”How Urban Concentration Affects Economic Growth”, *World Bank Working Papers*,n°2326.

[22] **Hsiao C.**, (1986), ”Analysis of Panel Data”, *Econometric Society Monographs*, n°11, Cambridge University Press.

[23] **Islam N.**,(1995), "Growth Empirics: A panel Data Approach", *Quarterly Journal of Economics*, 110(4), p.1127-1170.

[24] **Krugman P.R.**, (1991), *Geography and Trade*, Leuven University Press MIT Press.

[25] **Levine R., Loaysa N. and Beck T.**, (2000), "Financial intermediation and Growth: Causality and Causes", *Journal of Monetary Economics*, 46(1), p. 31-77.

[26] **Madariaga N., Montout S. and Ollivaud P.**, (2004), "Regional Convergence, Trade Liberalization and Agglomeration of Activities: An Analysis of NAFTA and MERCOSUR case", *mimeo.*

[27] **Mankiw N.G., Romer D. and Weil D.N.**, (1992), "A Contribution to the Empirics of Economic Growth", *The Quarterly Journal of Economics*, 107(2), p. 407-437.

[28] **Marina A.**, (2001), "Convergencia de los inbdicadores sociales en Argentina", *In Mancha Navarro and Sotelsek*, Ed. Pirámide, Madrid.

[29] **Martin P. and Ottaviano G.I.P**., (2001), "Growth and Agglomeration", *International Economic Review*, 42(4), p. 947-968.

[30] **Martin P. and Ottaviano G.I.P**., (1999), "Growing Locations : Industry Location in a Model of endogenous Growth", *European Economic Review*, 43(2), p. 281-302.

[31] **Moran P.A.P**., (1950), "A Test for Serial Independence of Residuals", *Biometrika*, 37, p. 178-181.

[32] **Niebuhr A.**, (2003), "Spatial Interaction and Regional Unemployment in Europe", *European Journal of Spatial development*,p.1-26.

[33] **Porto G.**, (1996),"Las Economías regionales en la Argentina" , *GEL*, Buenos Aires.

[34] **Porto G.**, (1994),"Convergencia y Política económica. Algunos resultatdos para provincias argentinas" , textitAnales de la XXIXa Reunión Anual de la Asociación Argentina de Economia Política.

[35] **Rey S.J. and Montouri B.D.**, (1999),"U.S. Regional Income Convergence: A Spatial Econometric Perspective", *Regional Studies*, 33, p.145-156.

[36] **Rodriguez F. and Rodrik D.**, (1999), "Trade Policy and Economic Growth: A Skeptic's Guide to Cross-National Evidence", *NBER Working Papers*,n° 7081.

[37] **Slaughter M.J.**, (1999),"Per Capita Income Convergence and the Role of International Trade", *American Economic Review*, 87, p.194-199.

[38] **Upton G.J.G and Fingleton B.**, (1985),*Spatial Data Analysis by Example, Volume 1*, Wiley, Chichester.

[39] **Utrera G.E. and Koroch J.A.**, (1998), "Convergencia: evidencia empirica para las provincias argentinas (1953-1994)", *Anales de la XXXIIIa Reunión Anual de la Asociación Argentina de Economia Poltica.*

[40] **Vayá E., López-Bazo E., Moreno R. and Surinach J.**, (2000), "Growth and Externalities across Economies: An Emprirical Analysis using Spatial Econometrics", *Working Paper Economic and Social Sciences*, n°EOO/59, University of Barcelona.

[41] **Venables A.** , (1996), "Equilibrium Locations of Vertically Linked Industries", *International Economic Review*, vol.37, n°2, pp.341-359.

[42] **Willington M.**, (1998), "Un an álisis empírico del crecimiento económico regional en Argentina", *Estudios*,XXI(84), p 13-36.

[43] **Windmeijer, F.,** (2000), "A Finite Sample correction for the variance of linear two-step GMM estimators", *Institute for Fiscal Studies Paper 00/19*.

[44] **Yudong Y. and Weeks M.**, (2003), "Provincial Conditional Income Convergence in China, 1953-1997: A Panel Data Approach", *Econometric Reviews*, 22(1), p.59-77.